

## **Understanding Social Psychology across Cultures: Living and Working with Others in a Changing World**

Peter B. Smith, Michael Harris Bond and Cigdem Kağitçibasi

### **Chapter 2: Improving the Validity of Cross-Cultural Psychology**

Version of 16 May, 8,674 words

Our focus in this volume is primarily but not exclusively on what is typically considered to constitute social psychology. Indeed, since a main element in our perspective is that one cannot understand any particular behavior without taking account of the context within which it occurs, one could argue that all of psychology is (or at any rate, should be) social psychology. However, some aspects of behavior are much more context-dependent than others, and it is these that demand our closest scrutiny. In this chapter, we explore the ways in which the present understanding of how to do cross-cultural psychology has been achieved. This undertaking will entail a certain amount of historical review and a focus on methods of investigation. We must defer detailed consideration of theories and results until we have some confidence in the skills and techniques available to cross-cultural investigators. Throughout the chapter, we shall underline a series of methodological cautions that can be drawn from earlier experience.

Published accounts of the origins of psychological research are often written in a way that suggests a linear development toward a particular contemporary research method. Those who favor experimental research highlight the establishment of laboratories, for instance by Wundt in Leipzig, Germany in 1879. Those who currently favor qualitative or discursive methods point to the earlier foundation of a social psychology journal in Germany by

Steinthal and Lazarus (Lück, 1987) and to descriptive accounts of crowd behavior in France (Le Bon, 1895). English language authors typically identify the social psychology texts by McDougall (1908) in the UK and Ross (1908) in the USA as the first to be published, failing to note earlier texts by Tarde (1898) in French and Orano (1902) in Italian. In a similar way, Triplett (1898) working in the USA is often credited with having conducted the first empirical study of social influence processes, whereas Ringelmann had made studies on the same topic in France ten years earlier (Kravitz & Martin, 1986), as had Binet and Henri (1894). Even earlier discussions of how best to study social behavior empirically were reported from Italy (Cattaneo, 1864). It may seem pedantic to point out these studies, which now have only archival significance. However, they serve to emphasize our first caution:

**Guideline 1. Cross-cultural psychologists should not expect that all relevant studies will be published in English.**

Clearly there can be no definitive history of psychology, but rather a set of histories within which different authors highlight events that mark the way toward their own preferred present (Lunt, 2003). Our own account has a similarly selective perspective. While the work of most of those noted above implicitly assumes that human behavior can be validly studied by sampling at a single location, we look instead for work that tests for universality by using some kind of comparative research design. Three rather separate paths toward this perspective can be discerned. One path leads us to social anthropology, one comprises a psychometric perspective, and the last has emerged from a primarily North American social psychology. We consider these in turn.

## **A SOCIAL-ANTHROPOLOGICAL PERSPECTIVE BASED ON FIELDWORK**

Accounts of the focus of Wundt's psychological laboratory and of the persons who visited him make it clear that the current specialist focus of the different social sciences did not then exist. He was a skilled generalist. In his later publications, Wundt espoused what he termed *Völkerpsychologie*, and his visitors included those who subsequently became well known as sociologists and social anthropologists, for instance Durkheim and Malinowski. Wundt's account of *Völkerpsychologie* was not based on observations, but was focused on analysis of such issues as language, myth, customs, religion and art. Although Wundt never drew on this material, through the first half of the twentieth century the developing discipline of social anthropology yielded a rich harvest of ethnographic studies of non-Western societies. Initially, it was seen as worthwhile simply to document the practices and social relationships characterizing a given society. As more and more studies were published, it was inevitable that attempts would be made to construct theories as to why societies differed in the ways that they did.

Thus was born the 'culture and personality' school. The US anthropologist, Ruth Benedict, for instance characterized whole cultures in terms such as 'paranoid', which might be applied to individuals and which were derived from integrating observations about the typical individuals of that culture. As she put it:

'It is recognized the organization of the total personality is crucial in the understanding or even in the mere description of the individual personality. If this is true in individual psychology where individual differentiation must be limited always by the short span of human lifetime, it is even more imperative in social psychology, where the limitations of time and of conformity are transcended. The degree of integration that may be attained is of course incomparably greater than can ever be found in individual psychology. Cultures from this point of view are individual psychology thrown large

upon the screen, given gigantic proportions and a long time span.’ (Benedict, 1932, p.24)

Drawing on this perspective, Benedict (1946) asserted that the national character of the Japanese was based on shame, in contrast to that of Western nations that was said to be more focused on guilt. She acknowledged that some individuals might not fit the overall pattern, but was primarily concerned with understanding the overall profile of a given society.

Psychological anthropologists, as they became known, were strongly influenced by psychoanalytic theories (Piker, 1998). Once the massive databank of the Human Relations Area Files had been assembled, it became possible to test hypotheses linking particular styles of child rearing with the predominant practices of different societies. For instance, support was found for the hypothesis derived from psychoanalysis that societies in which weaning occurs relatively late would favor remedies for illness that are taken orally (Whiting & Child, 1953). The culture and personality school found particular applicability to societies that were small and homogeneous.

When we try to comprehend the social behavior of those who live in contemporary societies, it becomes imperative to make a distinction between the experiences of the individual and the cultural context within which they live. There is simply too much variability across the characteristics of individuals within the same cultural group to justify the simplifications arising from the culture and personality tradition. We discuss the reasons for making this distinction more fully in Chapter 7, but for the moment we can simply note a second methodological guideline arising from the historical development of the field:

**Guideline 2. Cross-cultural psychologists need a clear understanding of the distinction between individual-level analysis and culture-level analysis.**

An additional perspective drawn from fieldwork has been that which derives from the work of Vygotsky (1934/1962) in Russia. Vygotsky studied peasant communities in Central Asia, focusing upon the development of thought and language. He emphasized the manner in which learning derives from the socio-cultural context within which the child develops. For instance, the child acquires language skills from the more competent practitioners with whom he or she interacts. These models provide illustrations of ways of speaking and thinking that are slightly more complex than those the child has yet mastered, but still within its 'zone of proximal development'. The child first learns to speak in the new ways observed from his or her socialization models, and then internalizes these skills. This focus on cultural transmission of cognitive development of children has been important precursor to contemporary **cultural psychology** (Cole, 1996; Valsiner, 1989; Rogoff, 1990; Segall, Dasen, Berry, & Poortinga, 1999; see also the recommended Further Reading by Valsiner). The focus of cultural psychology is no longer distinctively upon children, but highlights the reciprocal process by which everyday interactions between individuals transmit and reformulate our sense of what goes on around us.

## **THE PSYCHOMETRIC PERSPECTIVE**

The basis upon which psychology became distinguished from neighboring social sciences was through its emphasis upon studying samples of individuals within controlled settings, rather than focusing upon larger groups, organizations or nations. Experimentalists test the specific effects of changes within a controlled environment on the individuals in that environment, but there is an equally strongly developed set of procedures based simply upon psychometric testing of individuals. If the beginnings of psychological study are marked by the use of these types of methods, then the beginnings of cross-cultural psychology are to be

found within the studies conducted by members of the Torres Straits expedition. Over 100 years ago, a group of social scientists visited the islands in the Torres Straits, which separate Australia and New Guinea. One member of the team, Rivers, focused his studies upon the islanders' perceptual processes. Rivers referred to his respondents as 'savages', and was influenced by a belief prevalent at that time that because 'savages' were less intelligent than Caucasians they might have superior visual skills:

'The natives were told that some people had said that the black man could see and hear better than the white man, and that we had come to find out how clever they were...'  
(Rivers, 1901).

Despite whatever preconceptions Rivers may have held, he obtained some rather striking results. For instance he reported that susceptibility to visual illusions varied, depending on the perceiver's **ethnicity**. The islanders were indeed less susceptible to the Müller-Lyer illusion, namely that the vertical lines with outward and inward facing arrow heads in Box 2-1 differ in length. However, they were more susceptible than Caucasians to the illusion that the vertical line in Box 2-1 is longer than the horizontal line. Susceptibility to various types of visual illusions has been studied extensively by cross-cultural psychologists in more recent times (Segall, Dasen, Berry, & Poortinga, 1999). The results are not wholly consistent, but it appears that differential sensitivity to illusions is a product of the particular type of environment in which one lives. The presence or absence of particular types of routine visual stimuli in the environment is thought to give differential encouragement to the development of relevant types of perceptual discrimination. This line of research raises two further issues of key relevance to cross-cultural psychology. Firstly,

**Guideline 3: If we wish to detect cross-cultural differences, tests are required that will validly detect those differences.**

- Box 2-1 about here –

Rivers used a great variety of tests, not just those that have been mentioned here. He was perhaps lucky, in that he chose to include two different tests of visual illusion. Had he used only one, he might have been led toward a false and possibly racist conclusion about the greater susceptibility of one group to illusions. Secondly, in explaining his experimental procedures to his respondents, Rivers chose to deceive them. The tests must indeed have seemed strange to the islanders, and some form of explanation was clearly required. In more recent times, psychologists also sometimes tell lies to their research participants, in the hope that this deception will conceal from them the hypotheses that they wish to test and enable the psychologist to assess their responses more validly.

The particular issue that informing participants about one's research raises for cross-cultural psychologists is this: if research participants are told that they are part of a cross-national comparison, is this likely to cause them to respond differently? In the case of the Torres Straits studies of perceptual processes, the answer is probably no, but in relation to social behaviors the answer is less clear. What would happen, for example, if a cross-cultural social psychologist told his or her participants that the purpose of the research was to compare national levels of helping behavior? This challenge leads to **Guideline 4: In conducting cross-cultural studies, care must be taken to explain one's study in ways that do not prejudice the validity of the data that are obtained.** Achieving validity does not require deception, but it may require an explanation that is only partial, if one is given at all.

Psychological research that involves testing of individual respondents' cognitive capacities has been very extensive in more recent times. Various types of intelligence tests have been frequently employed. Van de Vijver (1997) reported a meta-analysis that included fully 251 cross-national comparisons of cognitive abilities. **Meta-analysis** is a procedure that makes it possible to summarize studies using a wide range of different measures, by comparing the magnitude of the effects obtained, rather than by analyzing the results separately for each of the specific measures that were used (See also the Further Reading by Van Hemert for more on meta-analysis). Differences across nations in average effect sizes on measures of cognitive ability were most strongly affected by the age of respondents and by the affluence of the nations sampled.

In terms of Guideline 3 formulated above, the key issue in examining these results is to consider whether or not the various measures that were used can validly measure national differences in cognitive abilities such as intelligence. Some researchers used Western measures of ability, whereas others used locally developed measures. Van de Vijver did not, however, find a significant difference in the effect size between those who used Western measures and those who did not. For more social outcomes, the origin of the measure may make a difference.

There has been much controversy over the findings that certain ethnic groups and persons from certain nations score higher on intelligence tests than do others. It therefore matters a great deal whether 'culture-fair' measures have been used. The results of studies concerning visual illusions suggest that we might expect the types of intelligence that are fostered in different locations would depend on how life is lived in those locations. In other words, intelligence would need to be defined in terms of the abilities necessary to process information effectively that is relevant to the local context, rather than as an absolute quality of which one has either more or less. Western-style schooling may develop a focus on

particular modes of logical, abstract thought. Other contexts may require other ways of thinking. Mundy-Castle (1974) interviewed West Africans as to the qualities that they perceived to make up intelligence and found that they considered social skills an equally important aspect of intelligence to what he called technological skills. Segall et al. (1999) note further studies that reach similar conclusions. Rather than treating cognitive ability as a unitary trait, it is preferable to study the way in which different environments encourage the development of the kinds of cognitive ability that are needed for survival in each context (Sternberg & Grigorenko, 2003). Although the tests of cognitive ability that are employed in most cross-national surveys continue to emphasize the Western focus on abstract reasoning, it is interesting to note the recent surge of interest in Western nations in ‘emotional intelligence’ (Goleman, 1996). This development suggests that what was first identified as social intelligence in Africa actually has a relevance that is not restricted to effective functioning in less technically advanced societies.

- Box 2-2 about here –

The dilemmas faced by those seeking to measure cognitive abilities can be considered in light of a useful distinction proposed by Pike (1967) and popularized by Berry (1969). In attempting to learn a Mexican Indian language, Pike found that the use of differing pitches and tones influenced the meaning of specific sounds in that language. In terms of the concepts used in linguistics, phonetic production affects the meaning of specific phonemes. Drawing on this distinction, Berry contrasted two approaches to cross-cultural study. Firstly, one could start from the assumption that there are universals and proceed in that manner until evidence is found for differences. He termed this the ‘**etic**’ approach, paralleling the universalist assumptions made in phonetics, the study of sounds.

Alternatively, one can start by studying intensively the distinctive attributes of one specific cultural group. He termed this the ‘**emic**’ approach, because it focuses on local meanings, and draws most readily on information provided by persons within that cultural group. This orientation parallels linguists’ focus on the phonemic attributes of a specific language.

Berry suggests that most cross-cultural research is initially ‘**imposed etic**’, that is to say, it is based on Western concepts and measures applied in non-Western contexts. The assumption is made that the concepts and measures will have the same meaning in new contexts. The global use of intelligence tests such as the Wechsler Intelligence Scale for Children (WISC) and Wechsler Adult Intelligence scale (WAIS) is a typical example. Emic studies such as that by Mundy-Castle (1974) can highlight the limitations of using imposed-etic measures. As a research field becomes more fully developed, an accumulation of emic studies can contribute to the development of improved ‘derived-etic’ measures, that have equal validity in a broad range of contexts.

A simple example of progress toward this goal can be found in cross-cultural studies of emotion recognition, which we shall discuss in Chapter 6. In early research, persons from many nations were asked to identify emotions portrayed in posed photographs of American faces. This procedure assumes that all possible types of emotional expression are to be found in US faces. In later studies, faces from all participating nations were included. Berry’s concepts lead us toward **Guideline 5: Cross-cultural researchers need to ensure that their stimulus materials and measures are understood comparably in each location.**

The creation of measures that are understood equally well and in similar ways across different parts of the world is not simply a matter of using items that refer to issues or tasks that are familiar to respondents. There is also a need to ensure that translations from one language to another are done in a manner that yields items with equivalent meaning. The most widely accepted procedure for achieving this is back-translation (Brislin, Lonner, &

Thorndike, 1973). That is to say, a translation is first made from the language in which the test was originally developed in to the language of the society in which it is to be used. A second bilingual person is then asked to translate the items back into the original language, without having seen the original version. Comparison of the retranslated version and the original can then be used to detect problematic translations and to create an improved version through discussion between the two translators.

This discussion often focuses on the relative merits of a literal translation versus ‘**decentering**’. A decentered translation is one that does not use terms that have precise linguistic equivalence, but which draw on the cultural knowledge of the translators to use phrases that have equivalent meaning in the two languages. For instance, while English speakers discussing some misfortune might seek hope by claiming that ‘every cloud has a silver lining’, speakers of Mandarin Chinese would claim that ‘every cloud has a pink edge’. A de-centered translation would drop the specific descriptors in favor of a similar, more general saying, like ‘something good comes from any misfortune’. Some of the problems of creating a measure that is understood in a similar way in a whole series of locations are illustrated in Box 2-3.

- Box 2-3 about here –

Once a satisfactory translation of items to be included in a questionnaire has been accomplished, it is next necessary to determine whether responses to the separate items defining a scale correlate together in the same way as they did in the original language. Procedures that have been used to establish the adequacy of a measure in its country of origin must be repeated in the new locations where it is to be employed. Does the translated scale have adequate **reliability**, in other words a pattern of consistent responses to items that

measure the same quality? Reliability may have been lost through the translation process. Furthermore, do responses to different scales differ in the same way that **factor analysis** in the country of origin had shown? Finally, is there evidence such as correlations with other measures that supports the **validity** of the measures in the new context? **Guideline 6: Researchers require evidence that their measures are reliable and valid in each new cultural setting.**

## **THE EXPERIMENTAL PERSPECTIVE**

While the fieldwork and the psychometric perspectives have both contributed substantially to the development of cross-cultural psychology, the most distinctive approach characterizing much psychological research has long been an emphasis upon some form of experimental method. Psychologists have favored the experimental method because it offers the best chance for determining causal relationships between variables. Simplified settings are required if one is to set up an adequately controlled experiment, but this simplification has mostly been seen as a price worth paying for the development of a truly scientific psychology, one that can establish causal relationships.

Whilst the earliest applications of experimental method were not particularly influential, many of the key figures in the foundation of modern social psychology were those who did favor experimentation, particularly those who practiced in the US. The political turbulence of the 1930s led to prominent researchers such as Kurt Lewin and Muzafer Sherif seeking refuge in the United States. Other key figures including Fritz Heider and Solomon Asch had migrated to the US a little earlier. Furthermore, although the origins of social psychology were European, as a consequence of the widespread destruction and dislocation surrounding World War 2, the practice of social psychology was for a time

largely confined to the US (Farr, 1996). Lewin, Sherif and Asch all espoused theories that emphasized the effect of the immediate social context on behavior. The generation of US researchers that succeeded them sought ways to illustrate these effects of context experimentally. In order to do this, they created simplified settings in which experimental subjects were exposed to various kinds of social pressures. Over the years, these procedures and the theories that they were designed to test have become more sophisticated, but they all seek to explore the variable impact of the surrounding social field on those persons interacting in that setting.

A series of stages can be identified through which North American experimental social psychologists and those from other parts of the world have sought understanding of the similarities and differences between the outcomes of their studies. Initially, US researchers initiated collaborative work with researchers from other parts of the world, attempting to replicate the results that they had obtained back home. As we shall see, these studies often yielded problematic results. As social psychology became practiced more widely around the world, a second stage became apparent. Various critiques of experimental methodology were formulated. In some nations, researchers argued in favor of the development of '**indigenous**' psychologies, abandoning experimental method and using more culturally appropriate methods. In other nations, especially in Europe, the critique was focused on improving the experimental method, rather than on abandoning it. The third and current stage in this process is one where social psychologists from different nations are increasingly collaborating on an equal basis, and draw on theories and methods that are explicitly formulated to explain cultural differences. We consider these three stages in turn.

### ***Stage 1: Replications***

Replication of research studies is a crucial element in the establishment of their validity. Even within a single nation, it is not always the case that the results of a study would prove replicable, since there are many ways in which one sample of participants in a study might differ from one another. Probably the most widely replicated experimental study in social psychology has been the Asch (1956) conformity study. Asch showed that when a group of experimental accomplices repeatedly all give incorrect judgments as to which of several lines matched another line, naïve experimental participants rather often also gave incorrect responses. Asch interpreted these effects in terms of conformity, but we should note that only 38% of the responses made by naïve respondents actually conformity. The study could also be interpreted in terms of the independence of the remaining 62% of responses.

R. Bond and Smith (1996) reported a **meta-analysis** of 134 published Asch conformity effects. Of these effects, 97 had been obtained with US respondents, while the remainder was drawn from 16 other nations. Bond and Smith used the US data to estimate the effect of variations in the types of experimental procedures that were used, the type of respondents, and the data of the study. They were then able to discount these sources of variance when examining the amount of influence that occurred within the studies done outside North America. As Box 2-4 shows, the degree of group influence on conformity responses was less within Europe than it had been in the US, but it was greater in the rest of the world than it had been in the US.

- Box 2-4 about here –

The results of this extensive analysis pose two questions. Firstly, it appears that a standard experimental procedure produces different results in different parts of the world,

even after all variations that could be detected within a great number of US replications have been discounted. We need a theory to explain why this change in effect size might occur.

Bond and Smith tested various relevant culture-based explanations and we shall return to this study after cultural theories have been introduced in Chapter 3.

Secondly, it is necessary to consider what to call the specific experimental effect that Asch obtained. Asch called it conformity, and subsequent writers follow this usage.

However, in differing cultural contexts, it is possible that the same behavior may have different meanings. One possible way to investigate the meanings of a behavior is by asking those who engage in that behavior why they are doing it. Asch did so and obtained a variety of answers from his US respondents. Some said that they thought their eyesight must be defective; others wanted to avoid embarrassment by giving “wrong” answers publicly. We lack similar interview data from other nations. One possibility is that respondents may have given incorrect answers not to save themselves from embarrassment, but to save the others from embarrassment! Confronted with a group of people who were obviously giving wrong answers, someone who valued tact or sensitivity might choose to reduce the others’ humiliation by tactfully also giving wrong answers.

Other very well known US studies have also been quite often repeated outside the US, yielding results consistent with the Asch findings. For instance, Milgram’s (1961) study of obedience to a destructive authority figure was repeated in eight other countries (reviewed in Smith & Bond, 1998, Chapter 2). The levels of obedience that were found varied, but the effect was obtained in all studies. However, when many other studies have been repeated in new locations, the results are sometimes non-significant and sometimes even in the opposite direction from those obtained in the US. Amir and Sharon (1987) attempted to estimate the magnitude of the problem of non-replicability. Rather than focus on famous studies, they selected six studies that had been published in major US social psychology journals and

attempted to replicate their results, using both high school and university students within Israel. They deliberately selected studies whose design and methods would be appropriate to Israeli respondents. The six original studies had yielded 37 significant effects. Amir and Sharon succeeded in replicating only 16 of these outcomes within both their Israeli samples, and nine more within just one of their samples. They also found 27 significant effects that had not been found in the original studies. Thus it appears that the replicability problem is very substantial, even when assessed in two cultural systems that many would regard as relatively similar.

A particularly striking instance comes from studies of the so-called social loafing effect. Studies have shown that, in the US, individuals put less effort into a task when working with others than when working alone. Karau and Williams (1993) reported a **meta-analysis** of 147 social loafing effects obtained within the US and 15 obtained in Pacific Asian nations. On simple tasks, such as clapping one's hands or shouting, social loafing effects were obtained equally in Pacific Asia and the US. On more complex tasks, the Pacific Asian studies showed a complete reversal: people worked harder when they were in groups than when they were alone. Subsequent studies by Earley (1993) provided further insight into these results. Earley studied managers working on tasks in USA, China and Israel. Both the Israeli and the Chinese managers worked harder when they believed that their tasks were part of a group effort, whereas the Americans worked harder when they believed that they were working alone. In addition, some of those participants who thought that they were working in a group were led to believe that the group comprised others known to them, whereas other participants were told that they were working with strangers. The enhanced social effort expended by Israelis and Chinese was found only when working with one's own group, not with strangers. This study begins to provide clues as to the nature of the cultural

differences that may explain the reversal across nations of the social loafing effect, and will be explored further after we have considered the concept of **collectivism** in Chapter 3.

One of the requirements of experimental method is that one needs a supply of participants who can present themselves at a psychology laboratory without too much difficulty. This consideration provides one reason why experimental social psychologists have often based their sampling on students. Students differ from the general population in many respects, even within a single nation. However, there are additional hazards in basing cross-cultural comparisons based on student populations: The populations of students differ greatly across different countries. In Western nations, university education is undertaken by a relatively large percentage of the age cohort. However, in many nations participation in university education is achieved only by a small percentage of the age cohort, and is drawn disproportionately from elite families. Comparing student samples may give misleading results. Using other populations, as was done in Earley's studies, is preferable. **Guideline 7: Sampled populations should be comparable.**

### ***Stage 2: Decentering Research Methods***

We noted above that translators of research materials often need to use **decentering** to make a translation meaningful in another language. Likewise, the results of early attempts to replicate US studies in other locations indicated a need to decenter the research methods that were to be used in studies conducted in other parts of the world. Participation in a social psychology experiment requires one to treat as 'real' a staged event in which one typically interacts for a short period of time with one or more strangers. Valid responses to such a setting are much more likely within contexts where people habitually meet many strangers than in settings where they do not.

In developing their own studies, researchers from different parts of the world have put emphasis on the importance of what they saw as distinctive or '**indigenous**' perspectives. Markus and Kitayama (2003), for example, have concluded that, 'Psychology as we knew it was not yet a comprehensive psychology; it appeared to be the indigenous psychology of America or perhaps, more specifically, the psychology of secular, middle-class Anglo America.' (p. 280) By examining other indigenous perspectives and the results they produce, we can form an opinion as to whether different methods suit different national contexts, or whether the debate about methods transcends national boundaries.

We first consider the case of Europe. The initial recovery of social psychology in Europe after the Second World War owed much to visitors from the US (Van Strien, 1997). Reflecting on numerous early US visitors to his research group at Oxford in the UK, Argyle wrote:

'...all were a great source of stimulation, information and help. Our group became an important channel for the transfer of American social psychology to Britain. And yet we kept our distance from American social psychology. They had colonized us, perhaps intentionally, but we altered the message. We were impressed by their ingenious and well designed experiments, but we found them too artificial, insufficiently related to real behavior. We could not see how this kind of research could be applied to real problems. We were looking for a different way of doing it. The way we favored could also be found in several places in the US, but not in the mainstream' (Argyle, 2001, pp. 340-341).

Argyle's concern to address real problems and avoid artificial contexts was echoed by some of the early work in the US. For instance, the series of field experiments of

intergroup conflict conducted by Sherif and his colleagues each lasted several weeks (Sherif, 1966). However, these were exceptional. As European social psychology developed during the 1960s, critical perspectives were formulated of the way in which the experimental methods that were being used neglected the social context within which studies were being conducted (Israel & Tajfel, 1972). Tajfel, a leading proponent of this critique, argued that experiments were conducted 'in a vacuum', focusing only on the individuals and their response to a temporarily contrived, unrealistic social context. He argued for a greater role of small and large group affiliations in determining social behaviors. This challenge led to his formulation of social identity theory and its subsequent development into self-categorization theory (Tajfel & Turner, 1979; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). These theories have proved distinctively popular among European social psychologists. Researchers in this field continue to use experimental method, but they test predictions that emphasize the impact of group affiliations in real-world settings.

A second and more radical critique of experimental method in social psychology was articulated by the French social psychologist, Moscovici (1972). Moscovici developed a viewpoint that has become known as the social representations perspective. In a way that recalls the work of Vygotsky, he argues that social knowledge does not reside within individuals, but in communally shared representations of the persons, ideas and objects within one's environment. Researchers who study social representations have mostly abandoned experimental methods, and rely more on interviews and observation to elucidate the way in which one or other concept is socially represented. For instance, specific studies examine social representations of health, illness, the European Union, or the individual. Work on social representations is widespread in Europe and in some regions of Latin America. Both social identity theory and the social representations approach can be thought

of as having an indigenous European origin, but both now find adherents in other parts of the world (Smith, 2005).

Most other attempts to develop indigenous psychologies have involved a radical rejection of the experimental method and an attempt to develop methods appropriate to psychological study within a single nation (Kim & Berry, 1993; Sinha, 1997). The most extensive work of this kind is currently to be found in Taiwan and in the Philippines (Church & Katigbak, 2002). *Sykolohiyang pilipino* (Filipino psychology) entails distinctive methods of data collection built upon group participation and has yielded a set of indigenous concepts, among which *kapwa* is said to best describe Filipino social relationships. Enriquez (1993) cautions against inadequate translations of these concepts into English. He notes that while *kapwa* translates literally as ‘others’, the connotation of that word in English is one of exclusion, but in Tagalog *kapwa* implies inclusion with known, supportive others.

Part of the thrust of indigenous psychologists’ work is exemplified by the fact that little of it is available in English. Many Taiwanese studies are published in the Chinese language journal *Indigenous Psychological Research in Chinese Societies*. Yang provides a series of ten ‘Dos’ and seven ‘Don’ts’, arising from some of the Taiwanese work. Some of these are listed in Box 2-5. Note how Yang chooses to describe Western psychology as an indigenous psychology, not as mainstream psychology or as universal psychology.

- Box 2-5 about here –

Translations of concepts said to be indigenous do not always confirm their uniqueness, however. For instance Diaz Guerrero (1994) identified a series of more than 100 ‘historic socio-cultural premises’ that are widely endorsed by Mexican respondents and which form the basis of the indigenous Mexican ethnopsychology that he and his colleagues

have formulated. These premises comprise attitude statements in favor of what Diaz Guerrero terms affiliative obedience, *machismo*, respect over love and virginity. However, Ayçiçeği (1993) translated a selection of these statements into Turkish and found that Turkish respondents also endorsed them all. At various points in this book, we shall discuss several other indigenous concepts that are said to have meanings that do not apply to other cultures. There are likely to be many more that are not yet represented in the psychological literature and are fertile grounds for readers of this text to explore within their own cultural systems for possible overlap.

For the moment, it is more important that we focus upon the methodological dilemma that is raised increasingly by the fruits of indigenous research. How can we know whether concepts drawn from different languages do or do not have the same meanings? In the case of *kapwa*, there is a self-evident difference. Are such differences a major threat to the ultimate development of a unified psychology, or are the majority of differences a matter of nuance, which while important need not overshadow the presence of more general human universals? We can offer a provisional answer to these questions in the form of **Guideline 8: Evidence for universality is more likely to be found from a series of parallel studies done within different nations to explore the meanings of a concept, rather than from direct cross-national comparisons of mean scores on that concept.**

This is a more controversial guideline than those offered so far. Many cross-cultural studies involve direct comparisons of mean scores from one nation to another, as, for example, in the Asch ‘conformity’ studies discussed earlier. While studies of this type are important in identifying what appear to be cultural differences, they cannot tell us whether similar sounding concepts have the same meaning in different cultural contexts. To discover the meaning of a concept, we need to determine its correlates within each cultural context. The correlates of that concept define its ‘**nomological net**’, the set of meanings that make

sense of its usage and help social scientists build up a theory about the concept. If we find that a phenomenon has the same correlates in different cultural contexts, then we can be more confident that we are studying equivalent phenomena. As an example, let us revert to Earley's studies of social loafing. If we find that even though we get more loafing by groups in one nation and more loafing by individuals in another nation, the same predictors correlate with the presence or absence of loafing in both contexts, then we can be confident that we are comparing like with like. For instance, we might find that the level of endorsement of particular values in different nations could explain the change in loafing behaviors when alone or in a group.

### ***Stage 3: Benchmarks for Contemporary Cross-Cultural Studies***

In the preceding section, we gave some illustrations of the ways in which researchers in different parts of the world have expressed reservations about a psychology that is built upon global application of the experimental methods that are favored within mainstream US social psychology. Van de Vijver and Leung (1997) provide a much more comprehensive discussion of methodological issues in cross-cultural research (see also the Further Reading by Van de Vijver). Our position here is neither for nor against particular methods. Each can illuminate aspects that other methods may well miss. We have hinted already at some of the ways in which studies can be done that bridge at least some of the gaps between the divergent methods that are currently employed. Here we highlight some further achievements that have arisen from cross-cultural research over the past few decades, and which can be useful in evaluating the quality of the studies that we report in later chapters.

The first achievement to be noted is that recent studies are much more frequently based upon equal collaboration between those from inside and those from outside those

nations that are included in a cross-national study. Both internal and external perspectives are valuable in understanding patterns of social behavior. It is easy for **emic** researchers to think falsely that some aspect of their culture is unique; it is equally easy for **imposed etic** researchers to assume falsely that their measures capture all the important aspects of a particular culture. Together, they can provide a corrective to one another's blind spots. Equality is not always easily achieved, because some researchers may defer to others on the basis of seniority or presumed greater expertise. However, in the matter of internal versus external perspective, no one is more expert than anyone else. **Guideline 9: A better study will show evidence of input from both emic and etic perspectives.**

The second achievement concerns measurement. The two types of study that have been discussed in most detail in this chapter relied on objective measurements. In the Asch studies, social influence was measured by actual recorded judgments. In the social loafing studies, measures of actual work performance were used. However, very many of the studies reported by cross-cultural researchers utilize survey measurements of attitudes, beliefs, values or abstract judgments. All of these responses are recorded on rating scales, most typically those known as Likert scales, anchored by phrases such as 'Strongly agree' to 'Strongly disagree' or 'Very much' to 'Not at all'.

The study by Hofstede (1980) has been a key influence on cross-cultural psychology and will be discussed extensively in the next chapter. One particularly important aspect of Hofstede's study was that he identified the varying propensity of respondents in different parts of the world to record agreement or disagreement with Likert-scale items. Hofstede overcame this problem by the use of **within-subject standardization** of scores, sometimes known as ipsatization. In other words, he reasoned that given a full range of mutually contradictory statements, a respondent could not logically agree with all of them. Consequently, in order to make a valid comparison of mean scores across different nations, a

respondent's scores on each item must be made *relative* to his or her scores on all the other items. Hofstede made this type of correction for whole nations, but it can be done equally well for individuals. Studies will be discussed in later chapters that failed to make this type of correction and consequently drew conclusions about national differences that most likely reflected differences in **acquiescent response bias**, rather than the differences the researchers claimed to have studied. Box 2-6 illustrates the size of cultural differences that required correction in one recent study. **Guideline 10: A better study will be one in which the possibility of acquiescent bias has been taken into account, either by balancing items requiring positive and negative responses, or by estimating and discounting bias.**

- Box 2-6 about here –

The third achievement is that cross-cultural psychologists do now have theories to guide their studies. In the early years, many studies were conducted which made simple comparisons between a certain effect in country A and country B, without any initial rationale as to why these particular nations should have been selected for comparison. In the next chapter, we outline the emergence of the theories that are currently available to guide cross-cultural investigations. Progress toward this point has necessarily required debate as to how best to conceptualize culture and how to measure its salient attributes. Studies selected for discussion in later chapters will be evaluated in terms of all the guidelines outlined in this chapter, but with especial focus on this last one – **Guideline 11: The studies of greatest interest are those that test relevant theories about how culture affects the outcome studied.**

## SUMMARY

Studying psychological phenomena in differing cultural contexts requires skills additional to those required for research within a single population. This chapter has formulated a series of eleven guidelines for effective research that have arisen from the successes and failures of earlier studies. We reproduce them here for easy reference.

1. Cross-cultural psychologists should not expect that all relevant studies will be published in English.

2. Cross-cultural psychologists need a clear understanding of the distinction between individual-level analysis and culture-level analysis.

3. If we wish to detect cross-cultural differences, tests are required that will validly detect those differences.

4. In conducting cross-cultural studies, care must be taken to explain one's study in ways that does not prejudice the validity of the data that are obtained.

5. Cross-cultural researchers need to ensure that their measures are understood comparably in each location.

6. Researchers require evidence that their measures are reliable and valid in each new setting.

7. Sampled populations should be comparable.

8. Evidence for universality is more likely to be found from a series of parallel studies done within different nations to explore the meanings of a concept, rather than from direct cross-national comparisons of mean scores on that concept.

9. A better study will show evidence of input from both emic and etic perspectives.

10. A better study will be one in which the possibility of acquiescent bias has been taken into account, either by balancing items requiring positive and negative responses, or by estimating and discounting bias.

11. The studies of greatest interest are those that test relevant theories about how culture affects the outcome studied.

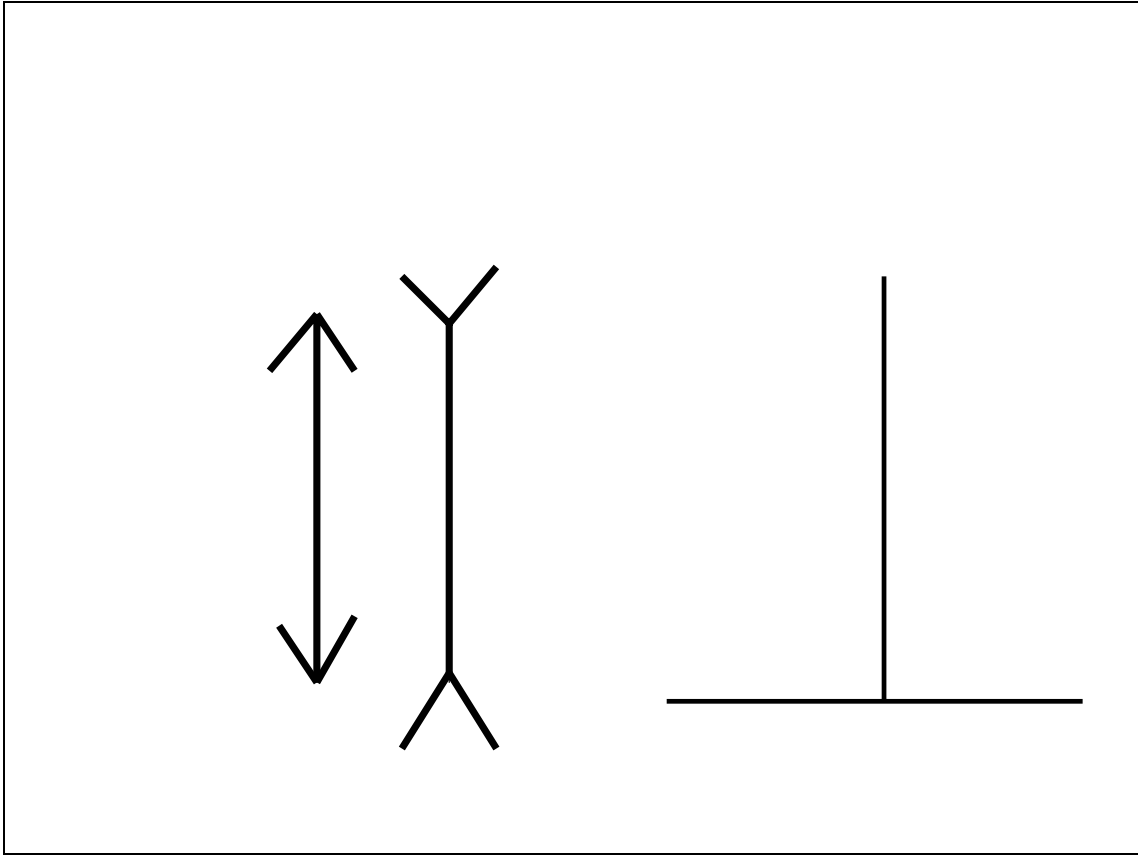
These guidelines can help us to evaluate the studies that are discussed in the rest of this book. Over the past century, some progress has been achieved towards designing and conducting studies that meet these criteria, but we have much more to learn before we can claim that we have a full understanding of culture's impact on social processes and behavior.

### **FURTHER READING**

1. Replications in other nations of early US studies are discussed in much fuller detail in Smith and Bond (1998, Chapter 2).
2. Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, 119, 111-137.
3. Valsiner, J. Culture and its transfer: Ways of creating general knowledge through the study of cultural particulars. [www.wvu.edu/psychology/~culture](http://www.wvu.edu/psychology/~culture) .
4. Van Hemert, D.A. Cross-cultural meta-analyses. [www.wvu.edu/psychology/~culture](http://www.wvu.edu/psychology/~culture)
5. Van de Vijver, F. J. R. (1997). Meta-analysis of cross-cultural comparisons of test performance. *Journal of Cross-Cultural Psychology*, 28, 678-709.
6. Van de Vijver, F.J.R. Types of cross-cultural studies in cross-cultural psychology. [www.wvu.edu/psychology/~culture](http://www.wvu.edu/psychology/~culture) .

### **STUDY QUESTIONS**

1. Should different research methods be used in studying different cultures or are there methods that can usefully be applied in all cultural contexts?
2. Which research method holds greatest promise for advancing our understanding of cross-cultural issues: fieldwork, psychometric tests and surveys, or experimentation?
3. Select any one of the 11 guidelines for cross-cultural research and explain why it is important in the study of social psychology across cultures.

**Box 2-1: Visual Illusions**

### Box 2-2: When is a Difference a Cultural Difference?

National differences in scores of intelligence tests focus on abilities considered without reference to context. What happens if we make comparisons that include context? After one year of schooling, children in Finland, Germany, Italy, Spain and Greece all achieve accuracy levels of greater than 90% on reading of words and of non-words constructed by the researcher. Children in Portugal and Denmark achieve around 70% accuracy. Children in the UK achieve an accuracy of just 40%. It takes them three or four years to achieve 90% accuracy.

Why is this very large difference found? Is it because children in the UK often start school between the ages of 4 and 5, while those in most other European countries start later? Is it because teachers in the UK are less well trained in the specific skills required to teach reading? Are UK teachers less motivated?

The principal explanation for the difference is that English is a much more difficult language to read. It is relatively easy for a Spanish or Greek child to learn to read, because their languages are phonically consistent. In other words, a given combination of characters always has the same sound. In English, vowels are frequently pronounced in different ways depending upon the consonants with which they are paired. Consider for instance the sound of 'a' in 'cap', 'call' and 'car', or of 'o' in 'go' and 'do'.

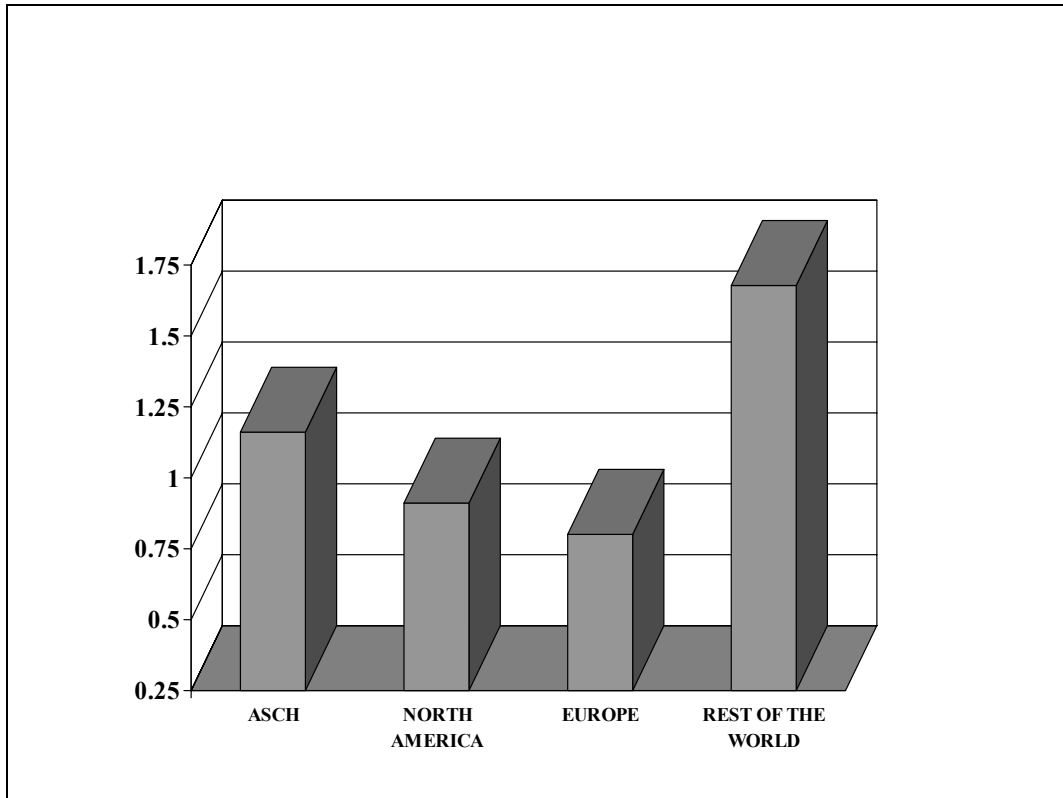
Comparisons of rates of learning to read differ from comparisons of intelligence scores, because languages differ from one another, whereas intelligence testers attempt to make their tests 'culture-fair'. Should differences in learning to read be considered as a cultural difference, or should differences only be attributed to culture when measures are used that are equivalent across cultures? Is language a part of culture? These questions are discussed in Chapter 3.

Sources: Seymour, Aro, & Erskine (2003); Goswami, Porpodas, & Wheelwright (1997).

### Box 2-3: Expressing Anger

Spielberger and his colleagues developed a measure of the different ways in which persons might handle their feelings of anger. Within the United States, he distinguished those who hold anger in, those who let anger out and those who are aware of their anger but actively control it. He used **factor analysis** to develop his State-Trait Anger Expression Inventory (STAXI). This measure has been translated subsequently into various languages and employed as an imposed-etic measure in studies including Canada, Norway, Austria, Italy and Singapore. However, it was found that there are fewer words in Norwegian describing the experience and expression of anger than in English. Some English metaphors concerning anger that are used in STAXI (e.g. burning up) have no direct equivalent in Norwegian; conversely some Norwegian metaphors 'Jeg r forbannet' ('I am cursed') have no direct equivalent in English. In Singapore, it was found that the 'anger-in' item "I am secretly quite critical of others" loaded on the 'anger-out' factor. This occurred because Singapore Chinese respondents interpreted the item as referring to talking negatively behind someone's back, while Western respondents interpreted the item in terms of holding grudges and not talking about them. In designing measures with cross-cultural validity for comparative purposes, items with variable meanings need to be identified and discarded.

Sources: Håseth (1996); Tanzer, Sim, & Spielberger (1996)

**Box 2-4: Average Effect Sizes found in Asch Replication Studies**

Source: R. Bond & Smith (1996)

**Box 2-5: How to do Indigenous Psychology**

**Do** tolerate ambiguous or vague states of understanding and suspend decisions as long as possible in dealing with theoretical, methodological and empirical problems, until something indigenous emerges in your mind during the research process.

**Do** be a typical native in the cultural sense when functioning as a researcher....

**Do** take the studied psychological or behavioral phenomenon and its sociocultural context into consideration.....

**Do** give priority to the study of culturally unique phenomena...

**Do** base your research on the intellectual tradition of your own culture rather than on that of a Western culture....

**Don't** neglect Western psychologists' important experiences in developing their own indigenous psychologies, which may be usefully transferred to the development of non-Western indigenous psychologies (emphasis added)

Source: Excerpts from Yang (2000)

**Box 2-6: Yea-saying and Nay-saying**

Business managers from 43 nations were surveyed. Respondents were asked to rate on 5-point scales to what extent they relied on each of eight sources of guidance in handling each of eight separate work events. Thus, each respondent provided 64 ratings, and it is not plausible that any manager would rely on all eight sources of guidance in handling all eight events. Below are shown the highest and lowest scores for what the researchers termed the ‘culture mean’ This is the mean across all 64 ratings – the net tendency to agree or disagree with statements in general. Note that the difference of 0.82 between the highest and lowest scoring nations spans 20% of the whole 5-point scale, indicating that **acquiescent response bias** could be a major source of misunderstanding in interpreting cross-national comparisons.

- Insert Figure from next page here –

In a further study, estimates of **acquiescent response bias** were derived from seven different surveys, each of which had sampled at least 34 nations. Substantial consensus across studies was found as to which nations tend to score higher and which tend to score lower in yea-saying. The nations that score highest are those in which respondents most strongly endorse collectivist family values and the need to reduce uncertainty.

Sources: Smith, Peterson, & Schwartz (2002); Smith (2004a)

